

Citation for published version:

Shaw, H, Ellis, D, Geyer, K, Davidson, B, Ziegler, F & Smith, A 2020, 'Quantifying smartphone 'use': Choice of measurement impacts relationships between 'usage' and health', *Technology, Mind, and Behavior*, vol. 1, no. 2. <https://doi.org/10.1037/tmb0000022>

DOI:

[10.1037/tmb0000022](https://doi.org/10.1037/tmb0000022)

Publication date:

2020

Document Version

Peer reviewed version

[Link to publication](#)

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Quantifying smartphone ‘use’: Choice of measurement impacts
relationships between ‘usage’ and health

Heather Shaw^{*1}, *David A. Ellis*², *Kristoffer Geyer*¹, *Brittany I. Davidson*^{2,3},
*Fenja V. Ziegler*⁴ and *Alice Smith*⁴

¹Department of Psychology, Lancaster University, Bailrigg, Lancaster, United Kingdom, LA1 4YW

²School of Management, University of Bath, Bath, United Kingdom, BA2 7AY,
<https://orcid.org/0000-0001-6172-3323>

³Department of Computer Science, University of Bristol, Bristol, United Kingdom, BS8 1QU

⁴School of Psychology, University of Lincoln, Brayford Pool, Lincoln, United Kingdom, LN6 7TS

*Heather Shaw (corresponding author), E-mail: h.shaw5@lancaster.ac.uk, <https://orcid.org/0000-0003-1449-3160>

Disclosures

Funding: This work was part funded by the Centre of Research and Evidence on Security Threats [ESRC Award: ES/N009614/1]

Conflict of interest: None of the authors have any financial, personal or organisational conflicts of interest.

Acknowledgements: Would like to thank Adam Birkinshaw and Foivos Vantzios for useful comments and suggestions throughout the project. Would also like to thank Neil Shaw, Lee Shaw, and Dr. Flora Ioannidou, for helping pilot the smartphone applications. We would also like to thank Dr. Jean-François Stich, Dr. Niklas Johannes and Dr. Linda Kaye for their useful comments and proposed revisions, which greatly improved the manuscript.

Open Science: All materials, data, and analysis scripts for this article can be found here: <https://osf.io/a4p78/>. Postprint is available here: <https://psyarxiv.com/mpxra>

Abstract

Problematic smartphone scales and duration estimates of use dominate research that considers the impact of smartphones on people and society. However, issues with conceptualisation and subsequent measurement can obscure genuine associations between technology use and health. Here, we consider whether different ways of measuring ‘smartphone use’, notably through problematic smartphone usage (PSU) scales, subjective estimates, or objective logs, leads to contrasting associations between mental and physical health. Across two samples including iPhone ($n=199$) and Android ($n=46$) users, we observed that measuring smartphone interactions with PSU scales produced larger associations between mental health when compared with subjective estimates or objective logs. Notably, the size of the relationship was fourfold in Study 1, and almost three times as large in Study 2 when relying on a smartphone ‘addiction’ scale instead of objective measures. Further, in regression models, only smartphone ‘addiction’ scores predicted mental health outcomes, whereas objective logs or estimates were not significant predictors. We conclude that addressing people’s appraisals including worries about their technology usage is likely to have greater mental health benefits than reducing their overall smartphone use. Reducing general smartphone use should therefore not be a priority for public health interventions at this time.

Keywords: Smartphones, Technology, Mental Health, Sedentary Behaviors, Screen Time, Methods, Digital Health

Introduction

Smartphones are primarily used for connecting people in a variety of personal and occupational settings. While the benefits of interpersonal communication are well-established (Berkman et al., 2000), most research concerning the relationship between communication, technology and health has focused on ‘negative consequences’ of smartphone use and screen time with a strong focus on mental health (Elhai, et al., 2017), and sedentary behaviours (Zagalaz-Sánchez, et al., 2019). Often referred to as ‘problematic smartphone use’ (PSU) or ‘smartphone addiction’ (Elhai et al., 2017), these refer to the perceived undesirable side-effects of use, which are mirrored in public discourse (Genc, 2014; Yang, Asbury, & Griffiths, 2019). However, there is a growing acknowledgement that the majority of research linking any screen time behaviours to health outcomes are themselves problematic (Science and Technology committee, UK Gov, 2019). For example, a growing number of academics have argued that research needs to address issues with measurement (Ellis, 2019), theory (Orben, 2018; Shaw, Ellis, & Ziegler, 2018), and analysis choices (Orben & Przybylski, 2019), by prioritising high-quality designs to better understand genuine benefits or harms (Coyne, et al., 2019; Heffer, et al., 2019). This may, in part, explain the lack of a coherent academic position regarding the impact of smartphone use on wellbeing. This remains troublesome when it comes to justifying the existence or effectiveness of interventions that aim to reduce usage. In this paper, we specifically investigate whether the relationship between smartphone use and health changes noticeably as a result of how smartphone use is conceptualised and measured.

Usage and Psychological Well-Being

Survey research has repeatedly linked increased smartphone screen time to lower psychological wellbeing (Twenge, Martin, & Campbell, 2018). However, many have noted that smartphone use is rarely measured directly, despite objective data being readily available from devices themselves (Ellis, et al., 2019; Twenge, 2019). Moreover, in recent years, concerns regarding ‘overuse’ have led to an abundance of usage scales being created to measure new constructs, including: ‘addiction’, ‘nomophobia’, and ‘problematic use’ (Ellis, 2019; Thomée, 2018). Specifically, when using problematic smartphone use scales, research consistently links higher scores with greater mental health symptomology, however these relationships seem to either dissipate or lessen when collecting duration estimates of use or objective logs (Elhai et al., 2017; Harwood et al. 2014; Rozgonjuk et al., 2018; Katevas, Arapakis, & Pielot, 2018; Vahedi & Saiphoo, 2018). Thus, understanding when and why these inconsistencies occur remains essential.

Usage and Physical Health

Beyond psychological impacts associated with usage, research has also linked greater smartphone use with increased sedentary behaviours (Lepp, et al., 2013; Zagalaz-Sánchez et al., 2019). Accordingly, people report that 87% of all phone use occurs while seated (Barkley & Lepp, 2016), and similarly, 90.9% of users report that they typically are sitting when using their smartphone (Xiang et al., 2020). Thus, it has been proposed that increased smartphone use lowers energy expenditure due to sedentary behaviours, and it is this mechanism, which results in greater body fat and higher rates of obesity (Hamilton, Hamilton, & Zderic, 2007; Kim, Kim, & Jee, 2015). However, while 9 out of 14 articles in a recent systematic review

showed a negative relationship between smartphone use and physical activity, none of the articles measured smartphone use objectively via logs from the device itself (Zagalaz-Sánchez et al., 2019). Instead, people self-reported the duration and frequency of their smartphone behaviours, which is widely documented to only have moderate correlations with actual usage (Andrews et al., 2015; Boase & Ling, 2013; Parslow, Hepworth, & McKinney, 2003; Ellis et al., 2019; Kobayashi & Boase, 2012; Lee, et al., 2017; Vrijheid et al., 2006). Therefore, research linking physical activity or sedentary behaviours to smartphone use is also scarce and yet to be examined precisely using objective logs.

Conceptualising Smartphone Usage

When understanding mental health relationships, more nuanced approaches suggest that how users think about and appraise their own smartphone usage is uniquely related to wellbeing and can be considered separately from objective use of the device itself. For example, a recent study found no evidence linking objective use of social applications to momentary wellbeing (Johannes et al., 2019). However, they did observe that the more positively people felt about their technology-mediated interactions in the past half hour, the better they felt in the current moment (Johannes et al., 2019). In addition, when assessing email use in occupational settings, stress levels increase when a person perceives their usage to be greater or lower than desired (Stich, et al., 2019). This suggests that people aim to regulate technology usage as they would with other everyday behaviours including for example, social affiliation (O'Connor and Rosenblood, 1996). Negative or positive appraisals may be dependent on whether a person has been able to achieve their preferred amount of usage (O'Connor and Rosenblood, 1996; Stich, et al., 2019). Thus, the way people perceive their

smartphone usage behaviours (e.g. a belief that their use is excessive) may drive relationships with mental health, that are independent from actual usage.

While there is no consensus regarding how smartphone usage or screen time should be conceptualised or measured, documenting ‘usage’ is of interest to many (Ellis, 2019).

Researchers however, continue to conflate the measurement of smartphone usage with assessing an individual’s appraisal of use. For example, defining or measuring problematic smartphone use (PSU) in relation to ‘overuse’ or ‘excessive use’ is prevalent in many articles (Elhai & Contractor, 2018; Elhai, et al, 2020; Kim, 2017; Yang et al., 2019). This has foundations in the Behavioural Addictions framework, where tolerance is a key component (e.g., the need to increase use over time to get the same ‘fix’) (Billieux, Maurage, et al, 2015; Elhai et al., 2017; Kim, 2017). Hence, it is not surprising to find questions such as “*Using my smartphone longer than I had intended*”, and “*Having tried time and again to shorten my smartphone use time but failing all the time*” in problematic usage scales (Kwon et al. 2013). However, agreeing with these statements only shows that a person is negatively appraising their smartphone use, and is not a measure of frequency or screen time in itself.

Correspondingly, research that has attempted to quantify the relationship between problematic usage scales and objective logs report many small effect sizes (Ellis et al., 2019), and exploratory factor analysis research shows that PSU scores do not cross-load with factors representing actual usage (Davidson, Shaw, & Ellis, 2020). This evidence already suggests that people’s appraisals of their smartphone use and actual usage should be considered separately.

In light of this unclear conceptualisation, it is important to distinguish between PSU as a psychological construct that appraises use, and smartphone usage as a behavioural variable,

because it has implications for theory and any proposed treatment. For example, if negative associations with physical and mental health are driven entirely by usage appraisals, then providing interventions that focus on usage behaviours alone may not deliver any benefits (Loid, Täht, & Rozgonjuk, 2020).

The Present Study

Measuring the associations between health and smartphone use in different ways could generate radically different results when relying on different operationalisations: subjective estimates, objective logs, and psychometric scales. This paper aims to understand this issue by collecting all three measures from the same participants. We therefore asked the question:

“Do problematic usage scale scores generate larger associations with health when compared with estimates of usage or objective behavior from the same users?”

Furthermore, we examined if increased smartphone use, when measured objectively, could account for variability in physical or mental health. Therefore, we also ask:

“Can objective smartphone use (pickups and screen time) account for differences in mental health symptomatology or physical health?”

These questions were first investigated during exploratory analysis of 46 adults who completed all three measurements, alongside an assessment of their body composition and anxiety, depression and stress symptomatology. The results were then used to generate hypotheses regarding the influence of different usage measurements on effect sizes. A second

study then acted as a replication and provided increased statistical power. All materials for both studies are located on the Open Science Framework (see Shaw et al., 2018).

Study 1

Methods

Participants

The sample consisted of 46 [12 male] participants that were staff and students from the University of Lincoln, UK. This deviates from our preregistered sample size of 84 due to laboratory access and technical issues. However, posterior calculations determined that a total sample size of 44 was adequate to investigate two-tailed medium-to-large effect sizes ($r > .4$) with a power of .8 when $\alpha = .05$. Age was skewed, as we tested predominately younger adults [$M = 23.54$, $SD = 8.25$]. All participants were Android smartphone users and stated they exercised less than 10 hours per week.

The study was advertised around a University campus using posters, leaflets, subject pool systems, and social media channels during term time and during public engagement events. Therefore, the sample consisted of those who emailed the researcher in response to these advertisements. Participants were told they would receive a graph of their phone use and a printout of their health analysis as incentives to take part. Those recruited through subject pool systems received course credit in compensation for their time.

Measures

Study 1 collected numerous variables to explore the relationships between individual differences and objective smartphone use. For brevity, the focus of this manuscript is to describe the body composition and mental health relationships with general smartphone use. Therefore, only the variables and data collection procedures related to this aim are described here. For further information on the additional variables collected, see *supplementary materials*.

Objective Smartphone Use

Objective smartphone data was collected using an application developed specifically for the project called Activity Logger (Geyer, 2018). This ran on Android devices and collected data to the resolution of one second. Activity logger was set up to listen to three events: the phone being turned on, the screen being activated, and the screen being turned off. Background operations then took this information, retrieved the current time stamp, and stored this in internal memory. This data file was then exported via the application and contained a list of records where a UNIX time stamp was paired with an event stating whether the screen was turning “ON” or “OFF”. Source code for the application is available to download (<https://osf.io/a4p78/>).

Estimates of Smartphone Use

To gather estimates of daily smartphone screen time, participants were asked one question: “Think back to days 2 - 8 of the study. On average, how many hours a day did you spend on

your smartphone?”. Participants responded in hours and minutes. To measure people’s estimates of how many times a day they ‘picked up’ their device, participants were asked: *“Think back to days 2 - 8 of the study. On average, how many individual times did you use your smartphone a day? Think of these as individual pick-ups.”*

Problematic Smartphone Use

Smartphone addiction was measured using the Smartphone Addiction Scale (SAS), which contained 33 items (Kwon et al., 2013). Participants rated the extent to which they agreed to several statements, for example *“Feeling pleasant or excited while using a smartphone”*. Participants responded on a six-point Likert-Scale ranging from “Strongly Agree” (1) and “Strongly Disagree” (6). Higher scores indicated greater addiction risk. This scale was chosen because it is widely cited and correlates highly with a variety of other PSU measures, which all appear to measure the same construct (Ellis et al. 2019; Thomée 2018; Davidson, Shaw & Ellis, 2020).

Anxiety

Symptoms of anxiety were measured using the GAD-7 (Spitzer, et al. 2006) and included 7 items. Participants were asked *“how often in the last two weeks have you been bothered by...”* and responded on a four-point scale whereby 0 = “Not at all” and 3 = “Several Days”. Using >10 as a cut-off point, the GAD-7 has been shown to have 89% sensitivity and 82% specificity with a diagnosis of general anxiety disorder (Kroenke et al., 2007).

Depression

Severity of depression was measured using the PHQ-9 (Kroenke, et al. 2001). Each of the nine questions related to a criterion mentioned in the DSM-IV for depression. Participants were asked “*how often in the last two weeks have you been bothered by...*” and responded on a four-point scale whereby 0 = “Not at all” and 3 = “Several Days”. Using >10 as a cut-off point, the PHQ-9 has been shown to have 88% sensitivity and 88% specificity with a diagnosis of major depression (Kroenke et al., 2001).

Perceived Stress

The Perceived Stress Scale (Cohen, Kamarck, & Mermelstein, 1983) had 14 items which measured ‘the degree to which situations in one’s life are appraised as stressful’. Participants responded how often they felt a certain way on a 5-point Likert scale whereby 0 = “Never” and 4 = “Very Often”. Participants were asked questions such as “*In the last month, how often have you felt that you were on top of things?*”. Higher scores indicated greater perceived stress.

Objective Health Measures

Height was measured using a meter stick, with age and gender captured via self-report questions. This data was inputted as controls in subsequent bioimpedance analysis. Body composition was measured using the eight electrode Tanita MC-780MA body composition monitor. This provided an estimate of a person’s body fat percentage, body mass index, and skeletal muscle mass percentage, using bioelectrical impedance measures. Bioelectrical

impedance assessment using the Tanita MC-780MA was a good alternate to Magnetic Resonance Imaging and Dual Energy X ray absorptiometry (DEXA) which are costly, and time-consuming (Verney, et al., 2015). Notably, the Tanita MC-780MA produces body fat assessments which highly correlate with DEXA assessment ($r = .85$) providing concurrent validity (Verney et al., 2015).

Procedure

The study lasted nine days (see Fig. S.1 for an infographic of the itinerary). On day one, a lab session provided participants with study information, including example data, followed by a consent form and an online questionnaire. Participants answered questions, including date of birth, gender, and psychometric scales that were beyond the scope of this manuscript (*see supplementary materials*). Once completed, participants were guided through the installation of the Activity Logger, and the researchers documented the smartphone brand and operating system. All screen savers were set to turn off after 30 seconds, and the application was 'white listed' in the smartphones' battery settings, ensuring that the phone would not limit the applications' functionality if the smartphone battery was low. Participants were then asked to keep their phone switched on for the duration of the study, and to not close the application. Whilst the application should re-start independently, as a precaution, if a participant's phone was switched off during the week, or they closed it, participants were instructed to re-open the application. Participants were then provided with information detailing how to prepare for the body composition assessment on day nine. To control for factors influencing body composition results, participants were asked to refrain from intense exercise and alcohol up to 12 hours prior to the assessment. They were also asked to remain hydrated and book a time

in the afternoon that was three hours after lunch. All participants were asked to go to the toilet before this session.

Participants were requested to use their phone as normal and carry on with their everyday activities across days two to eight of the study. This ensured that seven full days' worth of smartphone data was collected for each participant. On day nine, they returned to the lab and upon arrival, emailed data from the application to a researcher. Next, participants completed a questionnaire containing the stress, anxiety, depression, smartphone addiction scales, and other measures not reported in this manuscript (*see supplementary materials*). They were then asked to provide an estimate of how much they picked up their phone, and the amount of time they spent on their phone, on average each day, across days 2-8.

Height was measured for the bioimpedance assessment. Participants were instructed to remove any jewellery, items in pockets, metal accessories, and were then asked to stand bare foot on the Tanita MC-780MA body composition monitor while holding the hand electrodes by either side of their body, without touching their legs. A 0.5kg clothing allowance was inputted into the Tanita software if participants were wearing light clothing (gym gear), and a 1kg clothing allowance was inputted for heavy clothing (jumpers, jeans). Upon completion, participants were given a printout of their body composition, a graph of their application use, and a graph of their screen time across the week. Finally, participants were debriefed and thanked for their time.

All procedures received ethical clearance by the School of Psychology Research Ethics Committee at the University of Lincoln and complied with British Psychological Society Guidelines (British Psychological Society, 2018). In the debrief, participants were told that

the study would not offer any clinical diagnosis of any disorders and were provided with information about charities and services if they needed further support. The study also underwent a data protection plan. Participants had full control of their data as phone logs were stored solely on their devices and could be deleted by the participant at any point during the study by simply uninstalling the application.

Results

Data processing

Data and analysis scripts for study 1 can be found on the Open Science Framework (<https://osf.io/a4p78/>). The median daily hours-of-use was calculated across days two to eight for each person to remove the influence of any extreme “Screen On” events that occurred if the phone battery depleted and the application did not log a ‘Screen Off’ event. Daily pickups (frequency of use) were averaged across days two-eight, in accordance with recent work (see Ellis, et al, 2019). For the smartphone addiction scale, GAD-7, and PHQ-9, the responses were summed to create a total score for each scale. Specific questions within the perceived stress scale required reverse coding, and then an overall sum was created per person. See Table 1 for a list of the variables used in the analysis and their descriptives.

Exploratory Analysis¹

When collating all 46 participants' data together, smartphone use was highly skewed, as 54.44% of uses were under 30 seconds in duration, and 43.54% of uses were under 15 seconds in duration. Due to this skew, we followed Bishara and Hittner (2017) recommendations and conducted Spearman Rank order correlations with Fieller, Hartley and Pearsons (1957) variance when calculating 95 % confidence intervals as these are robust against non-normality. To explore how differences in smartphone measurement may influence associations with health, Spearman correlations were conducted between all the health and smartphone variables (see Table 2). Notably anxiety, depression, and, stress had significant positive correlations with smartphone addiction scores (all p 's <.01), which did not occur with any other smartphone measure (see Fig. 1 for objective screen time specifically). In terms of effect sizes, smartphone addiction scores generated $|r_s|$ equal to or larger than .39 with mental health variables, whereby estimates and objective variables were lower (all $|r_s|$ <.2) (see Table 2; Fig. 3).

¹ Our original design and analysis plans were preregistered: <https://osf.io/5g9v6>. However, following advice from reviewers, several deviations can be summarised as follows. First, when considering the impact of different smartphone usage metrics on health, our original design grounded key predictions using the displacement and goldilocks hypotheses. References to both of these have now been removed. Second, we only compared six well-being variables with smartphone use throughout rather than 18 collected as part of study 1 (see supplementary materials). This was to ensure that findings from study 1 could be confirmed and replicated in a larger sample (study 2). Third, we now report daily pickups (any smartphone use) instead of daily checks (uses under 15 seconds) to again ensure parity between the two studies. Finally, median daily screen time was calculated instead of average daily screen time to control for any long 'ON' durations. These could occur if a smartphone was unable to record an 'OFF' log until power is restored following battery depletion.

Table 1: Study 1 descriptives.

Health Variables	Mean	SD	α	Smartphone Variables	Mean	SD	α
Anxiety	6.13	5.56	.92	Median Daily Screen Time (hrs)	3.74	1.60	.90
Depression	6.57	5.25	.85	Average Daily Pickups	133.18	63.52	.93
Stress	24.61	8.42	.87	Daily Screen Time Estimate (hrs)	5.08	3.36	
Body Mass Index	24.84	5.86		Daily Pickups Estimate	48.74	39.96	
Body Fat %	26.97	8.86		Smartphone Addiction Scale	90.09	21.20	.90
Skeletal Muscle Mass %	41.35	6.40					

Table 2. Spearman correlations between smartphone and health variables from study 1.

Health Variable	Smartphone Addiction		Screen Time Estimate		Pickups Estimate		Median Daily Screen Time		Average Daily Pickups	
	Spearman r_s	Spearman 95% CI	Spearman r_s	Spearman 95% CI	Spearman r_s	Spearman 95% CI	Spearman r_s	Spearman 95% CI	Spearman r_s	Spearman 95% CI
Mental Health										
Anxiety	.44**	0.17, 0.66	.11	-0.19, 0.40	.05	-0.25, 0.34	-.00	-0.30, 0.30	.11	-0.20, 0.39
Depression	.39**	0.11, 0.62	.19	-0.11, 0.47	-.05	-0.35, 0.25	.05	-0.25, 0.34	.08	-0.23, 0.37
Stress	.53***	0.27, 0.71	.18	-0.13, 0.45	.03	-0.27, 0.32	.00	-0.30, 0.30	.03	-0.27, 0.32
Physical Health										
Body Mass Index	-.25	-0.51, 0.05	-.10	-0.39, 0.21	-.14	-0.42, 0.16	-.32*	-0.57, -0.03	-.39**	-0.62, -0.11
Body Fat %	.09	-0.21, 0.38	.18	-0.13, 0.45	-.01	-0.31, 0.29	-.01	-0.30, 0.29	-.12	-0.40, -0.18
Skeletal Muscle Mass %	-.06	-0.35, 0.24	-.14	-0.42, 0.17	.05	-0.25, 0.35	.06	-0.24, 0.35	.19	-0.11, 0.47

Notes: * significant at $p < .05$, ** significant at $p < .01$, *** significant at $p < .001$. Alpha's remain uncorrected for multiple comparisons.

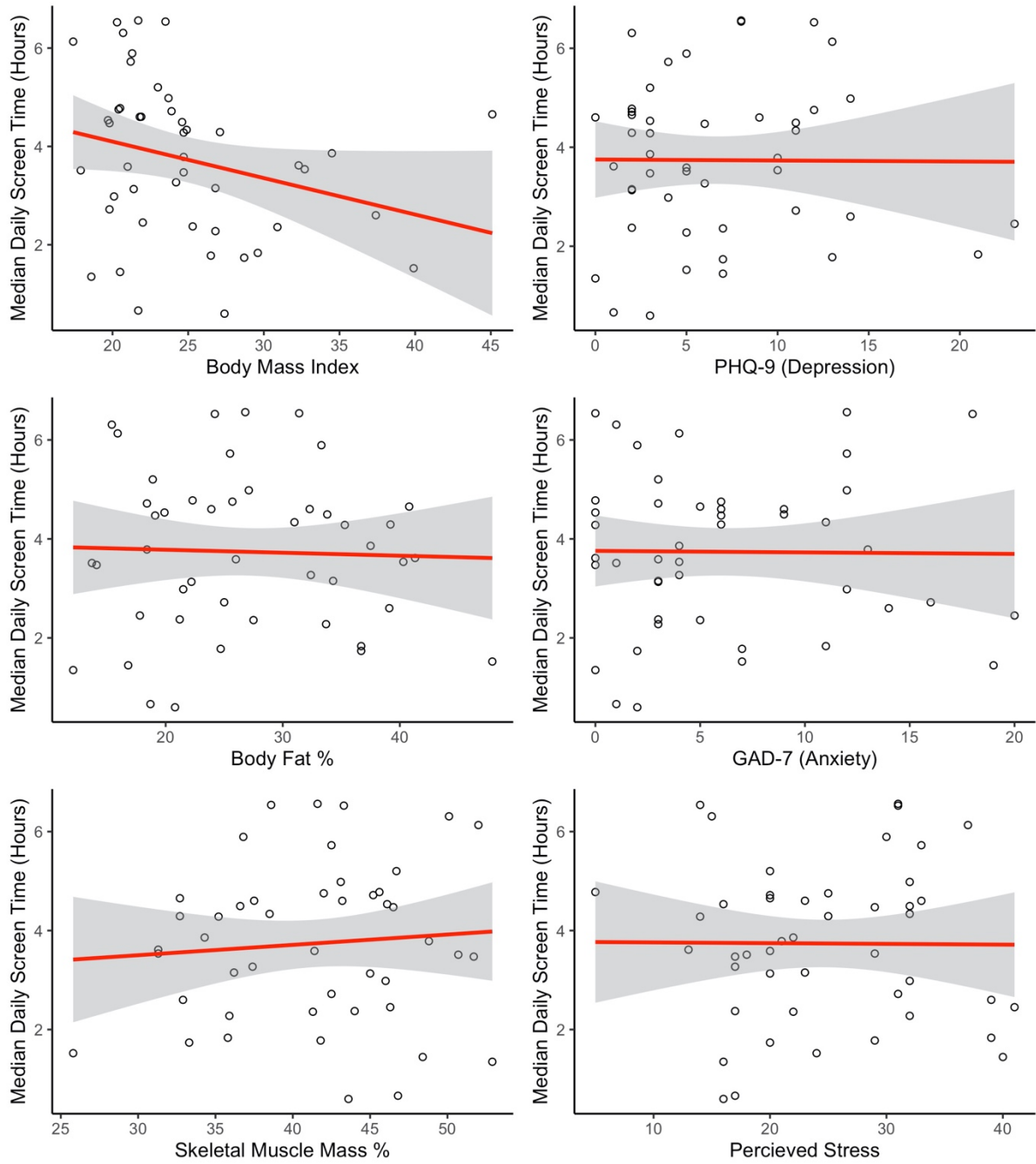


Figure 1. Scatter plots illustrating linear relationships between median daily screen time (Hours) and six health variables; body mass index ($r_s = -.32$), body fat percentage ($r_s = -.01$), skeletal muscle mass percentage ($r_s = .06$), depression ($r_s = .05$), anxiety ($r_s = .00$), and stress ($r_s = .00$). Regression line (red) illustrates linear relationship between each pair with 95% confidence interval (grey).

Discussion

In study 1, smartphone addiction was found to positively correlate with anxiety, depression, and stress measures. Pertinently, effect sizes quadrupled when measuring smartphone usage with a problematic usage (addiction) scale in comparison to objective screen time and pickup measures. In line with prior work, people's appraisals of their smartphone usage had stronger relationships with mental health than self-reported frequencies of use (Vahedi & Saiphoo, 2018) or objective logs (Rozgonjuk et al. 2018). This suggests peoples' appraisals of their smartphone use (e.g., worries) are more pertinent to mental health symptomatology than actual usage. Therefore, even within the same participants, a researcher could make different conclusions based on the measurement tool adopted. This is especially problematic when confounding the construct of problematic smartphone use with actual usage. Interestingly, we found that BMI reduced as daily screen time and pickups increased. While gravitating in the same direction, the effect size was smaller for correlations between actual usage and body fat percentage. Nevertheless, neither suggested the presence of any adverse effects between daily smartphone screen time and pickups on these measures of physical health.

We marked these findings as tentative until they could be replicated in a larger sample. This was examined in study 2, where we collected identical mental health and smartphone measures as study 1. We also re-assessed BMI and took advantage of retrospective data collected on a user's device, including daily logs of steps, and daily logs of 'walking and running' distance. Based on our previous findings, we predicted that effect sizes of $|r_s| > .3$ would be found when comparing mental health relationships with problematic usage scales,

and that lower effect sizes of $r_s^2 < .2$ would be found when examining estimates of use and objective logs².

Study 2

Methods

Participants

199 [137 women] participants, were recruited via Prolific Academic, from a subject pool of 24,117 iPhone owners. This pool contained predominately citizens from the United Kingdom and the United States. Participants had a mean age of 30.18 [$SD = 9.46$] and were paid £1.25 for their time. 42.71% of the sample were overweight or obese, and the average BMI across all participants was slightly higher than the recommended range [$M = 25.17$, $SD = 5.38$]. This was to be expected in a representative sample, as 52% of people have a BMI over 25 worldwide (WHO, 2018). A priori power calculation was performed which showed during two-tailed analysis a sample size of 192 participants was enough to detect small effect sizes of $r_s^2 \geq .2$ with a power of .8 when $\alpha = .05$.

²While preregistered analyses have been reported in full, our hypotheses primarily focus on effect sizes rather than just statistical significance following reviewer recommendations.

Measures and Procedure

Once clicking the link to access the online questionnaire, participants were presented with study information and a digital consent form. If participants agreed to take part, they were then asked; “*Please estimate how many hours and minutes you spend on your phone each day*” and answered in hours and minutes. In addition, participants were asked: “*Please estimate how many times a day you pick up and use your phone*”. After, smartphone addiction, anxiety, depression, and stress were then measured using the same scales as study 1.

Objective smartphone usage data was retrieved by utilising the Apple Screen Time feature that resides in modern iPhones. We used the same methodology as reported in Ellis et al. (2019) and extracted data retrospectively from the previous 7 days. In short, participants were prompted to find the ‘Screen Time’ graph and the ‘Pickups’ graph in Apple Screen Time settings and record for each day the number of pickups and screen time (in hours and minutes). For more details, see Ellis et al. (2019).

After obtaining objective smartphone use data, the questionnaire asked people to input their health data. The Apple Health App automatically tracks users’ steps, and their combined ‘walking and running’ distances. This historic data is accessible on a user’s iPhone for the entire time they have owned their iPhone. When clicking on the ‘Today’ tab, participants had access to a calendar where they could view their activity for any past day. Daily steps were collected by asking participants to click on the calendar pages for dates in the past week and enter for each day the number of steps displayed. Daily ‘walking and running’ distances were

collected by asking people to click on the calendar pages for dates in the past week and report the documented distance in either kilometres or miles. Participants were also asked if they owned a fitness tracker or a smartwatch and specified whether this device was synced to the Apple Health App. Lastly, participants were asked to report their age, gender, weight and height. They were given the option to answer in either metric (meters and centimetres / kilograms) or imperial measures (feet and inches / stones and pounds). At the end of the questionnaire, participants were debriefed, thanked for their time, and were then re-directed back to the Prolific Academic website.

All procedures received ethical clearance by the School of Psychology Research Ethics Committee at the University of Lincoln and complied with British Psychological Society ethical guidelines for internet mediated research (Hewson et al., 2013). Akin to study 1, the debrief provided websites where participants could access guidance regarding their mental health and were provided with details of 24-hour support lines. Participants could withdraw at any time before, during or up to two weeks after they completed the study by emailing the researcher.

Results

Data removal

Data and analysis scripts for study 2 can also be found on the Open Science Framework (<https://osf.io/a4p78/>). The survey received 263 respondents. However, this became 207 after removing those who did not have iOS12 installed, did not have an iPhone 5 or later, did not have seven days of screen time data on their smartphone, or did not complete the survey or

health questions. Another person was removed after being identified as an outlier when plotting data; they reported weight and BMI values more than three standard deviations from the mean. Finally, seven people were removed due to input errors (typos) in their health data. This left 199 participants for analysis. This was greater than the sample size derived from our preregistered priori power analysis (192).

Data coding and processes

Table 3 contains the descriptive statistics for all variables. Average daily screen time and average daily pickups scores were computed per person by taking the daily amount of screen time/pickups from the first six days and then calculating the mean. Six rather than seven days were used to compute this mean, as data from the seventh day did not represent a full day. Raw estimated number of daily pickups and estimated average daily screen time (in hours) were used in the analysis. Smartphone addiction, anxiety, stress and depression scales were all scored in the same way as study 1.

The daily physical activity variables; average daily steps and average daily ‘walking and running’ distance (km) were created by selecting the six days of data which corresponded to the same six days aggregated in the smartphone variables. The daily activity statistics from these six days were then averaged for each measure. If a participant reported their daily ‘walking and running’ distance in miles, this was converted to kilometres by multiplying the value by 1.60 before computing this average.

Lastly, BMI was calculated per person. Imperial height and weight responses were converted to metric units (centimetres and kilograms respectively). Finally, body mass index (BMI) was calculated from these values using the formula below:

$$BMI = Weight(kg) / Height(m)^2$$

Effect size analysis

Following study 1, to explore if differences in smartphone measurement influenced the size of the relationships with health, Spearman correlations were conducted between all the health and smartphone variables using Fieller, Hartley & Pearsons (1957) variance when calculating 95% confidence intervals (see Table 5). Spearman correlations were also conducted between all the smartphone measures to document differences between them (see Table 4; Fig. 2). Alpha's remain uncorrected for multiple comparisons.

Mirroring study 1, smartphone addiction scores consistently had effect sizes that were at least .36 or larger when correlated with mental health variables. Estimates and objective variables were lower (all $|r_s| \leq .21$) (see Fig. 3 or Table 5). This prompted an additional analysis that assessed whether this effect size deviation across measures was statistically significant. To compare differences in the magnitude between the coefficients, we adopted Hittner, May, and Silver's (2003) modification of Dunn and Clark's (1969) z test using the r package 'cocor' (Diedenhofen & Musch, 2015). This is suitable for the comparison of coefficients that are calculated from two dependent groups and share a variable in common (Diedenhofen & Musch, 2015). For example, it was possible using this method to compare whether the relationship between smartphone addiction and anxiety ($|r_s| = .43$) was statistically and

significantly larger than the relationship between average daily screen time and anxiety ($|r_s| = .16$). We also calculated Zou (2007) confidence intervals that rejects the null hypothesis if the interval does not include 0 (Diedenhofen & Musch, 2015; Zou, 2007). Findings showed that when assessing relationships with anxiety, depression and stress, that associations with smartphone addiction (PSU) were all significantly higher than the associations with estimates and objective logs (all p 's $< .05$) (see Table 6). The size of coefficients were not significantly different when using estimates or average daily screen time to determine associations with any mental health metric (all p 's $> .05$). However, there was a significant difference in effect sizes for mental health associations depending on whether an estimated or objective measure of pickups was employed, with correlations running in the opposite direction (all p 's $< .05$) (see Table 6 and Fig. 3).

Table 3: Study 2 descriptives.

Health Variables	Mean	SD	α	Smartphone Variables	Mean	SD	α
Anxiety	7.35	5.85	.94	Median Daily Screen Time (hrs)	4.62	2.30	.93
Depression	8.01	6.30	.90	Average Daily Pickups	85.76	39.94	.92
Stress	26.57	8.23	.85	Daily Screen Time Estimate (hrs)	4.38	2.15	
Body Mass Index	25.17	5.38	.25.17	Daily Pickups Estimate	47.14	39.81	
Average Daily Steps	5238.07	3345.92	.84	Smartphone Addiction Scale	105.80	24.36	.92
Average Daily 'Walking and Running' Distance	3.77	2.67	.83				

Table 4. Spearman correlations between all smartphone variables from study 2.

Smartphone Variable	Smartphone Addiction		Screen Time Estimate		Pickups Estimate		Average Daily Screen Time		Average Daily Pickups	
	Spearman r_s	Spearman 95% CI	Spearman r_s	Spearman 95% CI	Spearman r_s	Spearman 95% CI	Spearman r_s	Spearman 95% CI	Spearman r_s	Spearman 95% CI
Smartphone Addiction			.44***	.32, .55	.05	-.09, .19	.32***	.18, .44	.17*	.03, .31
Screen Time Estimate	.44***	.32, .55			.15*	.01, .29	.57***	.46, .66	.21**	.07, .34
Pickups Estimate	.05	-.10, .19	.15*	.01, .29			.10	-.04, .24	.30***	.16, .42
Average Daily Screen Time	.32***	.18, .44	.57***	.46, .66	.10	-.04, .24			.37***	.24, .49
Average Daily Pickups	.17*	.03, .31	.21**	.07, .34	.30***	.16, .42	.37***	.24, .49		

Notes: * significant at $p < .05$, ** significant at $p < .01$, *** significant at $p < .001$. Alpha's remain uncorrected for multiple comparisons.

Table 5. Spearman correlations between smartphone and health variables from study 2.

Health Variable	Smartphone Addiction		Screen Time Estimate		Pickups Estimate		Average Daily Screen Time		Average Daily Pickups	
	Spearman r_s	Spearman 95% <i>CI</i>	Spearman r_s	Spearman 95% <i>CI</i>	Spearman r_s	Spearman 95% <i>CI</i>	Spearman r_s	Spearman 95% <i>CI</i>	Spearman r_s	Spearman 95% <i>CI</i>
Mental Health										
Anxiety	.43***	0.31, 0.54	.21**	0.07, 0.35	-.08	-0.22, 0.07	.16*	0.01, 0.29	.16*	0.01, 0.29
Depression	.41***	0.28, 0.52	.19**	0.05, 0.32	-.10	-0.24, 0.05	.16*	0.01, 0.29	.17*	0.03, 0.31
Stress	.36***	0.23, 0.48	.21**	0.07, 0.34	-.10	-0.24, 0.04	.15*	0.01, 0.29	.12	-0.02, 0.26
Physical Health										
Body Mass Index	-.07	-0.21, 0.08	.09	-0.06, 0.23	.11	-0.03, 0.25	.16*	0.02, 0.30	.09	-0.5, 0.23
Average Daily Steps	-.16*	-0.30, -0.02	-.07	-0.21, 0.08	.26***	0.12, 0.39	-.07	-0.21, 0.08	.24***	0.10, 0.37
Average Daily 'Walking and Running' Distance	-.14*	-0.28, -0.00	-.07	-0.21, 0.08	.19**	0.05, 0.33	-.09	-0.23, 0.06	.17*	0.02, 0.30

Notes: * significant at $p < .05$, ** significant at $p < .01$, *** significant at $p < .001$. Alpha's remain uncorrected for multiple comparisons.

Table 6. Test's comparing differences in the magnitude of the coefficients when predicting mental health from varying smartphone variables. Each row in the table shows the z score when comparing variable 1's effect size with mental health to variable 2's effect size with mental health.

Variable one	Variable two	Anxiety		Depression		Stress	
		z	Zou's (2007) <i>CI</i>	z	Zou's (2007) <i>CI</i>	z	Zou's (2007) <i>CI</i>
Smartphone Addiction	Screen Time Estimate	3.14**	0.08, 0.36	3.11**	0.08, 0.36	2.10*	0.01, 0.29
Smartphone Addiction	Pickups Estimate	5.44***	0.33, 0.68	5.40***	0.33, 0.68	4.82***	0.28, 0.63
Smartphone Addiction	Average Daily Screen Time	3.48***	0.12, 0.42	3.20**	0.10, 0.40	2.65**	0.06, 0.36
Smartphone Addiction	Average Daily Pickups	3.16**	0.10, 0.43	2.80**	0.07, 0.40	2.74**	0.07, 0.41
Screen Time Estimate	Average Daily Screen Time	0.77	-0.08, 0.18	0.46	-0.10, 0.16	0.92	-0.07, 0.19
Pickups Estimate	Average Daily Pickups	-2.86**	-0.40, -0.08	-3.22**	-0.43, -0.11	-2.61**	-0.38, -0.06

Notes: * significant at $p < .05$, ** significant at $p < .01$, *** significant at $p < .001$.

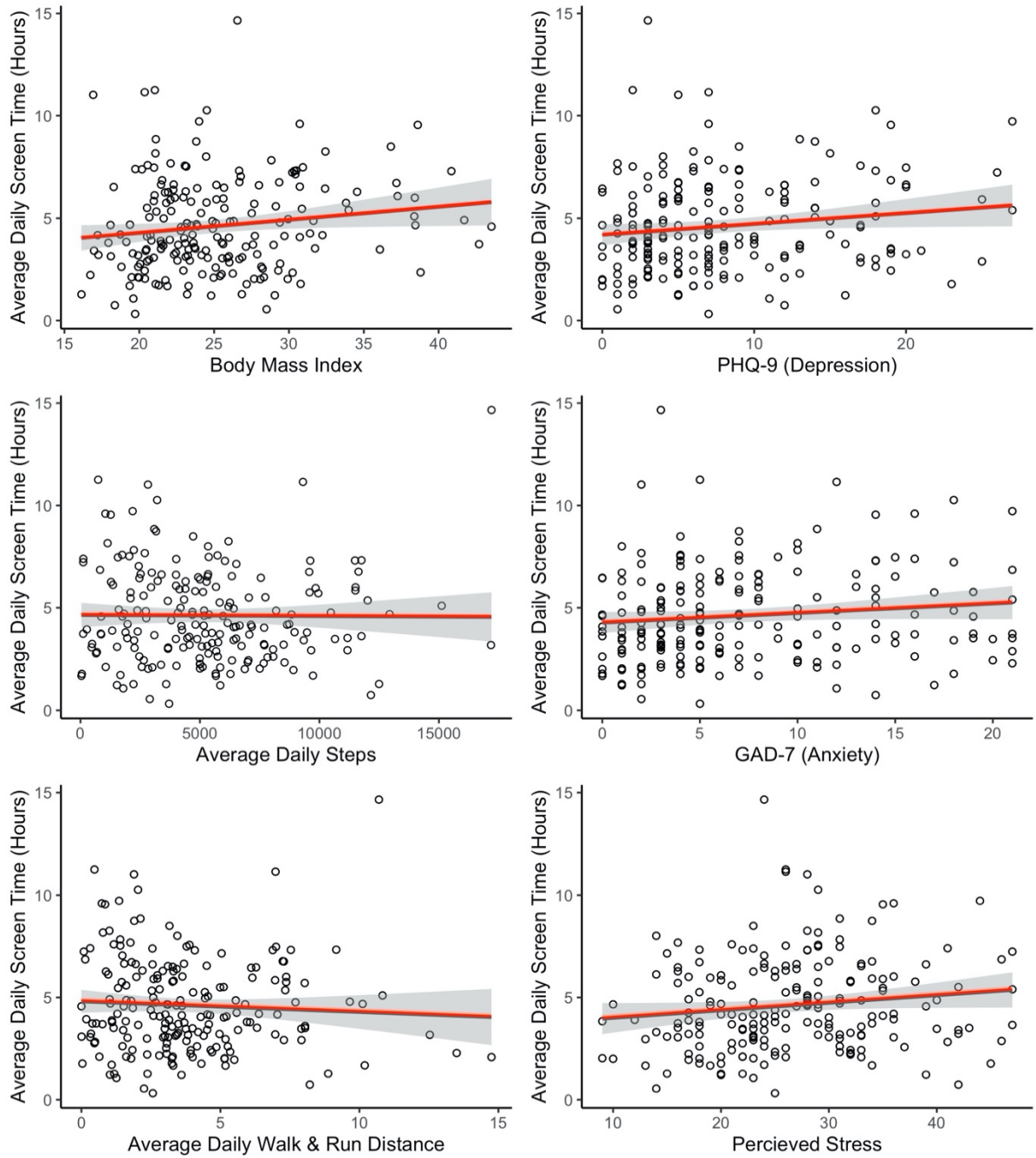


Figure 2. Scatter plots of linear associations between average daily screen time (hrs) with six health variables; body mass index ($r_s=.16$), averaged daily steps ($r_s=-.07$), average daily 'walking and running' distance ($r_s=-.09$), depression ($r_s=.16$), anxiety ($r_s=.16$), and stress ($r_s=.15$). Regression line (red) illustrates linear relationship between each pair with 95% confidence interval (grey).

Exploratory analysis – Tests of difference between groups with low and high mental health symptomatology.

Measuring ‘percentage variance explained’ through the exploration of effect sizes has been the subject of some criticism, with some authors advocating that significance testing between groups is a better indicator of whether screen time impacts mental health (e.g., Twenge, 2019). While this approach is in contradiction to many other statistical recommendations (Cumming, 2014), it was of interest to explore whether our conclusions would differ if we adopted this type of analysis. Consequently, as the GAD-7 and PHQ-9, have ‘cut off points’ (≥ 10) that indicate if people are at risk of having a disorder, we used these to create two groups; ‘low risk’ and ‘high risk’. These measures have high sensitivity and specificity (both $> .80$) when diagnosing depression and anxiety disorders (Kroenke et al., 2001; 2007). However, due to lack of further psychological assessment, we considered those who exceeded the defined cut-off points for each disorder to be at a higher risk, rather than define an individual as having the disorder. We then examined if people experienced different levels of daily smartphone use and PSU dependent based on group allocation.

To create groups for the analysis, 50 participants who were considered ‘high risk’ for both anxiety and depression were collated into one group. This group used their phone for an average of 4.72 hours a day ($SD = 2.27$) and picked up their phone on average 84.20 times a day ($SD = 37.98$). Those who didn’t exceed the cut-off values for either condition (scored less than 10 on both scales) were placed in a ‘low risk’ group ($n = 124$). This group used their phone for an average of 4.41 hours a day ($SD = 2.25$) and picked up their phone on average 84.07 times a day ($SD = 42.55$). Wilcoxon rank sum tests showed that the two groups did not significantly differ in their amounts of average daily screen time [$W = 3357, p = .39$] or

average daily pickups [$W = 3216, p = .70$]. This was mirrored when exploring differences in estimated daily screen time [$W = 3489.5, p = .19$] and estimated daily pickups [$W = 2721, p = .20$]. Therefore, those who were ‘high risk’ of having both general anxiety disorder and major depression did not use their smartphone’s differently to those who were ‘low risk’ for both conditions. However, a significant difference was found between the two groups on levels of smartphone addiction [$W = 4505.5, p < .001$]. Specifically, the ‘at risk’ group had higher smartphone addiction scores [$M = 116, SD = 23.67$] than the ‘low risk’ group [$M = 98.91, SD = 21.91$]. Consequently, if smartphone use is measured with subjective estimates or objective logs, we find no difference between ‘high risk’ and ‘low risk’ groups in terms of usage. However, if confounding usage and PSU, one would conclude the opposite if measuring ‘usage’, via a smartphone addiction scale, incorrectly positing that those with mental health symptomatology have higher usage.

Exploratory analysis – Linear Regression Models

Many researcher’s build predictive models to investigate if there is a linear or logarithmic relationship between health and smartphone usage (Csibi, et al., 2018; David, Roberts, & Christenson, 2018; Kim et al., 2016; Regan et al., 2020; Richardson, Hussain, & Griffiths, 2018). Following suit, we developed linear models that aimed to predict mental health symptomatology based on various smartphone variables. Notably, when including all five smartphone measures in models, only smartphone addiction scores significantly predicted mental health scores (see Table 7). Furthermore, models that only contained objective smartphone measures were not significant (all $R^2 \leq .02$, all p ’s $> .05$). Finally, average daily pickups significantly predicted average daily steps and average daily ‘walking and running’ distance across models (see Table 7).

Table 7. Linear regression models with health measures as dependent variables, and smartphone measures as predictors.

Model	<i>B. with criterion variable</i>						<i>B. with criterion variable (objective measures of usage only)</i>					
	Anx	Dep	Stress	BMI	Steps	Dist	Anx	Dep	Stress	BMI	Steps	Dist
Intercept	-2.95	-3.65	14.06***	25.05***	6061.84***	4.47***	5.46***	5.55***	24.15***	23.40***	4009.15***	3.35***
Average Daily Screen Time	-0.10	0.04	0.07	0.20	-36.04	-0.13	0.25	0.34	0.45	0.33	-129.31	-0.14
Average Daily Pickups	0.01	0.00	0.00	-0.00	20.20**	0.01*	0.01	0.01	0.00	0.00	21.29***	0.01*
Screen Time Estimate	0.14	-0.03	0.05	0.45*	-44.14	0.07						
Pickups Estimate	-0.01	-0.00	-0.01	0.02	10.01	0.00						
Smartphone Addiction	0.10***	0.11***	0.12***	-0.03	-25.22*	-0.02						
R^2	.18	.17	.13	.06	.10	.05	.02	.02	.02	.02	.06	.03
R^2_{Adj}	.16	.15	.11	.04	.08	.03	.01	.02	.01	.01	.05	.02

Notes: R^2_{Adj} = Adjusted R^2 B = beta estimates, * beta estimates significant to $p < .05$, ** beta estimates significant to $p < .01$, *** beta estimates significant to $p < .001$.

Anx = Anxiety, Dep = Depression, Stress = Stress, BMI = Body Mass Index, Steps = Average Daily Steps, Dist = Average Daily 'Walking and Running' Distance, All VIF scores between 1 – 2 and all tolerance scores $> .59$.

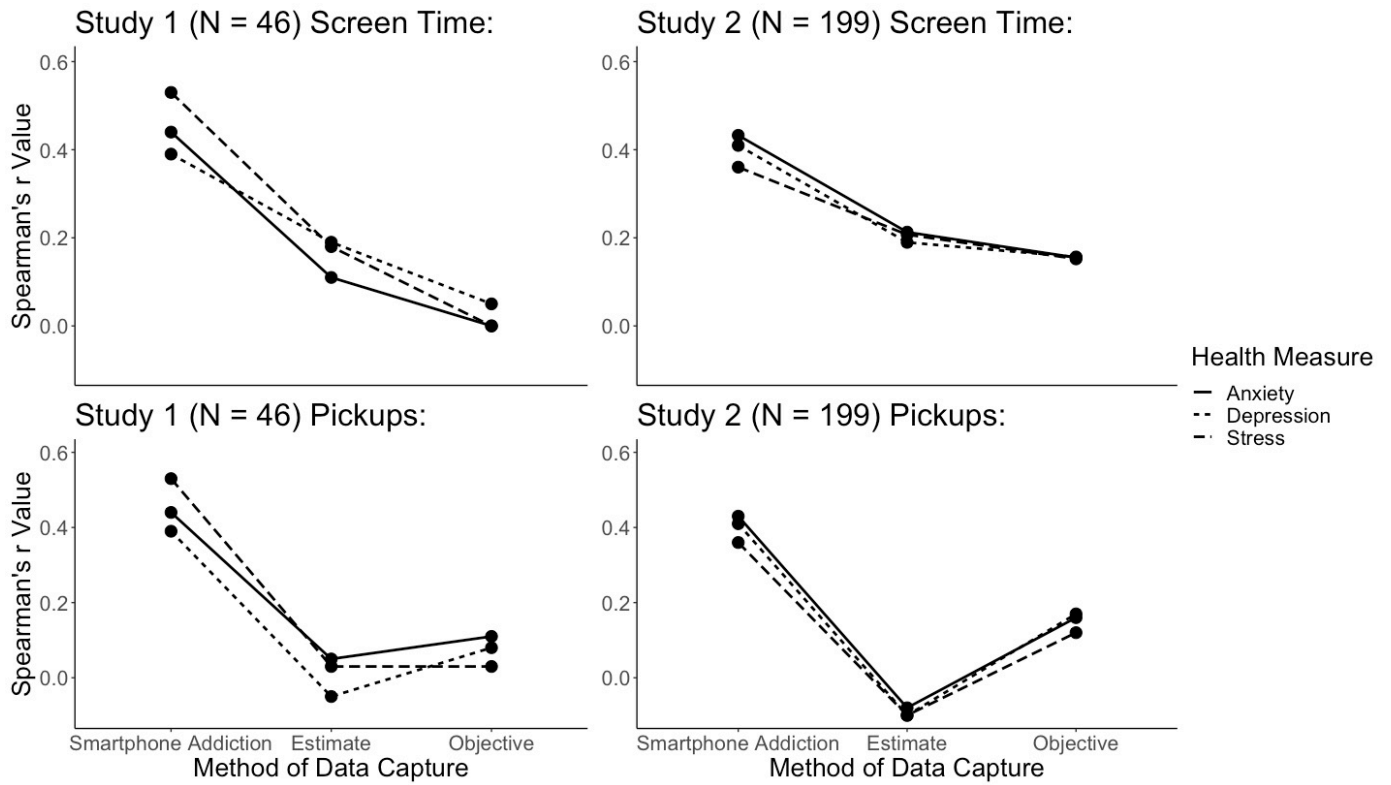


Figure 3: Visualising how a change in measurement effects relationships observed between smartphone use and depression, anxiety or stress across both studies. Top row illustrates how smartphone addiction scores, estimated and actual screen time correlate with mental health. Bottom row replaces estimated and actual screen time with pickups.

General Discussion

This paper considered if different conceptualisations and measurements pertaining to ‘smartphone use’, can generate contrasting associations with health. Across two samples including iPhone ($n=199$) and Android ($n=46$) users, we observed that PSU scales produced larger associations with mental health when compared with subjective estimates or objective logs. Notably, the size of the relationship was fourfold in study 1, and almost three times as large in study 2. Specifically, $|r_s| \leq .17$ were repeatedly found between objective smartphone use (daily pickups and screen time) and mental health symptomatology (anxiety, depression, and stress), whereas larger effects were observed when relying on a problematic usage scale (all $|r_s| \geq .36$). This was further supported with statistical models, which demonstrated that average daily pickups and average daily screen time do not significantly predict anxiety, depression, or stress, and explained less than 2% of the variance. Additionally, those who exceeded clinical ‘cut off points’ for both general anxiety and major depressive disorder did not use their phone significantly more than those who scored below a standard threshold. Finally, in terms of physical health, while previous research has observed associations between higher smartphone addiction scores and lower muscle mass (Kim, Kim, & Jee, 2015), our findings derived from objective logs are less clear-cut.

Generally speaking, conflating an individual’s appraisal of their smartphone use with actual usage generates vastly different relationships with well-being. This is problematic given a recent review confirmed that 70% of studies in this area adopt PSU scales (Thomee, 2018). The same review concluded that intense or frequent mobile use was associated with greater mental health symptomatology, yet this conclusion was based primarily on findings derived from PSU scales. Our findings alternatively suggest that helping people manage their

appraisals of use (e.g., worries) is more likely to provide a benefit to well-being than reducing use of the device itself. Consequently, one might question whether reducing actual smartphone use should be a priority for any intervention development at this time.

Recent research has arrived at broadly similar conclusions. For example, ‘intense’ general smartphone use did not predict negative wellbeing from objective logs (Katevas, Arapakis, & Pielot, 2018). Another study that measured objective smartphone screen time over a weeklong period, observed that average daily depressive mood positively correlated with smartphone addiction scores, yet objective screen time minutes were not related to depression and anxiety (Rozgonjuk et al. 2018). In terms of studies that rely on duration estimates, large-scale designs that follow Open Science practices have also reported small effect sizes. In a large sample of New Zealand adults ($n = 19,075$), associations between social media use and wellbeing were weak (Stronge et al., 2019). When using specification curve analysis to examine self-reports from a large sample of adolescents ($n = 355,358$), the association between digital technology use and wellbeing was again found to be small, explaining only 0.4% of the variance (Orben & Przybylski, 2019). In our sample, objective screen time and pickups explain less than 2% of the variance in mental health.

Placing our findings in a broader context, the relationship between objective use and mental health (all $r_s \leq .17$) is lower than the average effect size found across many psychology studies ($r = .21$). In comparison, this is slightly less than the relationship between Nicotine patch (vs. placebo) and smoking abstinence ($r = .18$), and about the same size as the relationship between post-high school grades and job performance ($r = .16$) (Funder & Ozer, 2019; Meyer et al., 2001). When adjusting for new recommendations that ‘small’, ‘typical’, and ‘relatively large’ effects fall around r coefficients of $\sim .10$, $\sim .20$ and $\sim .30$, respectively

(Gignac & Szodorai, 2016), the suggestion that social media has, for example, “*destroyed our lives*” would warrant moderate to large effects ($r > .20$) (Appel, Marker, & Gnambs, 2020, pp.62). Using this benchmark, our findings show that general smartphone use does not have extreme or profound effects on wellbeing, contrary to repeated claims suggesting otherwise (e.g., Twenge, 2017). At the same time, large effects of $r \geq .40$ in psychology studies are likely to overestimate a genuine effect and, as a result, warrant additional scepticism (Funder & Ozer, 2019). For example, the relationship between anxiety and smartphone addiction in study 2 was equivalent to the relationship between height and weight (both $r_s = .43$).

Scores from PSU scales may generate larger associations with mental health for several reasons. First, one could argue that negative appraisals of smartphone use (or technology use more generally) are based around issues that pertain to the regulation of everyday behaviour. Specifically, while people would like to perhaps regulate technology usage as they would with any other everyday behaviour, this is not always possible and this discrepancy between actual and desired use can lead to negative or positive appraisals (O’Connor and Rosenblood, 1996; Stich, et al., 2019). Second, both overall scores derived from the SAS and individual items have latent relationships with stress and depression scales (but not with objective smartphone measures) (Davidson, Shaw, & Ellis, 2020). Hence, cross-loadings between PSU and mental health could artificially inflate relationships due to a lack of independence. Third, ‘method bias’ may be influencing the size of correlation coefficients due to linguistic similarities between items across mental health and PSU scales (Podsakoff, MacKenzie, & Podsakoff, 2012). Every question in the SAS (and the majority of related scales) assesses a perceived problem, echoing mental health scales (Kwon et al. 2013; Spitzer, et al., 2006; Kroenke, et al., 2001). However, negative wording alone could be a further source of bias.

For example, it has been shown that correlations between role conflict, role ambiguity and other constructs reduced by 238% when controlling for wording effects, by balancing the number of positively and negatively slanted questions (Harris & Bladen, 1994; Podsakoff, MacKenzie, & Podsakoff, 2012).

Future Research

It is beyond the scope of this paper to discuss every issue pertaining to how technology use is conceptualised, measured, and analysed. However, future research that aims to specifically consider the impact of smartphone *use* should, where possible, adopt a more nuanced approach to understand both the costs and benefits of specific smartphone applications that can be monitored remotely (Geyer, et al., 2020). Recent work has shown that while total time spent using smartphones had $r = .16$ effect sizes with anxiety and depression (matching our work), certain categories of applications have beneficial relationships (e.g., time spent reading books) (David et al., 2018). Therefore, claiming general smartphone use as negative or positive oversimplifies a very complex and multifaceted phenomenon. For example, the relationships observed between body mass index and objective smartphone use were incoherent across our two studies. However, there appears to be a positive relationship between physical activity and objectively measured pickups. These results further question whether all smartphone behaviors should be considered sedentary when deliberating the relationship between usage and physical activity. Arora et al., (2013) found that computer use, tv viewing, and video gaming were associated with increased BMI, but conversely, did not find the same for mobile phone use. They stated, “*the portable nature of a mobile telephone does not require the user to remain in one place during use, thus allowing movement*” (Arora, et al. 2013, pp. 1258). In line with recent discussions, screen time is often

conceptualised without acknowledging ‘exergaming’ and other activities which involve physical activity whilst engaging with the device (Kaye, et al., 2020). Therefore, given that objective measures of technology use and exercise can be recorded by the same device, or in conjunction with a wearable tracker, future research should consider associations between specific patterns of usage and physical activity in greater detail. A variety of ecological momentary assessments including measures of cognitive functioning, mood or anxiety could extend these investigations further (Ellis, 2020).

We acknowledge that it remains difficult to objectively measure the use of a specific application across many devices (e.g., documenting time spent on Netflix across smartphones, televisions, and tablets) (Kaye et al., 2020), and researchers may still have to rely on estimates of use. However, our findings remain important as they confirm consistent discrepancies between objective logs and subjective estimates (see Table 4) (Andrews et al., 2015; Boase & Ling, 2013; Parslow, Hepworth, & McKinney, 2003; Ellis et al., 2019; Kobayashi & Boase, 2012; Lee, et al., 2017; Vrijheid et al., 2006). In study 2, and as observed previously, estimated frequency of ‘pickups’ had greater deviation from its objective counterpart than screen time estimates (Andrews, et al., 2015; Ellis, et al., 2019). Thus, if subjective estimates are to be collected, it is advised that researchers start including this measurement error into statistical models, which we have now quantified (Ellis, 2020; van Smeden, Lash and Groenwold, 2019).

Limitations

Both studies were cross-sectional; therefore, we cannot make any causal claims regarding the impact of smartphone use and mental health. However, by using a quasi-experimental

approach in the exploratory analysis of study 2 and through analysing the naturally occurring levels of mental health symptomatology in our sample, our findings cast doubt on the presence of any causal relationships that have been proposed previously, as those in a high symptomatology group did not have increased general smartphone usage. It is further possible that participants may have received feedback from Apple Screen Time prior to the study, which would have influenced their estimation of use. The size of the relationship between estimated screen time and actual screen time is larger in study 2 than previous work, and may explain why association between mental health and these two measures of usage did not significantly differ (Andrews, et al., 2015; Ellis, et al., 2019). However, this does not mitigate the need to control for errors between actual and self-reported screen time as part of any future analysis.

In addition, by moving our second study to an online platform, we achieved a larger and more representative sample. However, this meant losing some of the precision obtained with laboratory based bioimpedance measures when examining physical health. Nonetheless, as BMI scores in study 1 had large correlations with body fat percentage ($r_s = .70$) and skeletal muscle mass % ($r_s = -.73$), we accepted this as a relatively good proxy in study 2.

Furthermore, as self-reports of height and weight may also have measurement error, we analysed the ranges of BMI values. Our sample in Study 2 specifically had BMI values that were in line with what might be expected in representative sample (WHO, 2018). However, future research would benefit by exploring how body composition (including body fat percentage) could be collected objectively when relying on remote data collection.

Conclusions

To conclude, choosing between measurement tools, and accepting the benefits and limitations of that choice is an unavoidable facet of all research. However, when understanding or making claims regarding the effects of a particular behaviour on health, the cost of any error can be considerable. Here we demonstrate that problematic smartphone usage scales have significantly larger relationships with mental health when contrasted with objective logs of use. These are nearly thrice in a large sample and fourfold in a small sample. Thus, if a research question concerns technology usage, then objectivity should remain the preferred measure. The notion of ‘problematic use’ requires stringent examination because it is frequently conflated with behaviour despite a general acceptance that ‘excessive’ smartphone usage does not necessarily equate to ‘problematic use’ (Billieux, Philippot, et al., 2015; Panova & Carbonell, 2018). Consequently, PSU scales may only capture people’s appraisals of their smartphone use, rather than an underlying pathology or behaviour. Finally, our findings would favour addressing peoples’ appraisals about their usage rather than reducing their overall screen time, as the former relates more strongly to mental health symptomatology. Even if specific worries in relation to mobile technology are widespread, limiting general smartphone use or engaging with any form of ‘digital detox’ is unlikely to have any demonstrable benefits and should not be a priority for public health interventions at this time (Wilcockson, Osborne & Ellis, 2019)

References:

Andrews, S., Ellis, D. A., Shaw, H., & Piwek, L. (2015). Beyond Self-Report: Tools to Compare Estimated and Real-World Smartphone Use. *Plos One*, *10*, e0139004.

<https://doi.org/10.1371/journal.pone.0139004>

Appel, M., Marker, C., & Gnambs, T. (2020). Are Social Media Ruining Our Lives? A Review of Meta-Analytic Evidence. *Review of General Psychology*, *24*(1), 60–74.

<https://doi.org/10.1177/1089268019880891>

Arora, T., Hussain, S., Hubert Lam, K. B., Lily Yao, G., Neil Thomas, G., & Taheri, S. (2013). Exploring the complex pathways among specific types of technology, self-reported sleep duration and body mass index in UK adolescents. *International Journal of Obesity*, *37*(9), 1254–1260. <https://doi.org/10.1038/ijo.2012.209>

Barkley, J. E., & Lepp, A. (2016). Mobile phone use among college students is a sedentary leisure behavior which may interfere with exercise. *Computers in Human Behavior*, *56*, 29–33. <https://doi.org/10.1016/j.chb.2015.11.001>

Berkman, L. F., Glass, T., Brissette, I., & Seeman, T. E. (2000). From social integration to health: Durkheim in the new millennium. *Social Science & Medicine*, *51*(6), 843–857.

[https://doi.org/10.1016/S0277-9536\(00\)00065-4](https://doi.org/10.1016/S0277-9536(00)00065-4)

Billieux, J., Maurage, P., Lopez-Fernandez, O., Kuss, D. J., & Griffiths, M. D. (2015). Can Disordered Mobile Phone Use Be Considered a Behavioral Addiction? An Update on Current

Evidence and a Comprehensive Model for Future Research. *Current Addiction Reports*, 2, 156–162. <https://doi.org/10.1007/s40429-015-0054-y>

Billieux, J., Philippot, P., Schmid, C., Maurage, P., De Mol, J., & Van der Linden, M. (2015). Is Dysfunctional Use of the Mobile Phone a Behavioural Addiction? Confronting Symptom-Based Versus Process-Based Approaches. *Clinical Psychology and Psychotherapy*, 22(5), 460–468. <https://doi.org/10.1002/cpp.1910>

Bishara, A. J., & Hittner, J. B. (2017). Confidence intervals for correlations when data are not normal. *Behavior Research Methods*, 49(1), 294–309. <https://doi.org/10.3758/s13428-016-0702-8>

Boase, J., & Ling, R. (2013). Measuring Mobile Phone Use: Self-Report Versus Log Data. *Journal of Computer-Mediated Communication*, 18(4), 508–519. <https://doi.org/10.1111/jcc4.12021>

British Psychological Society (2018). *Code of Ethics and Conduct (2018)*. Retrieved May 09, 2020 from: <https://www.bps.org.uk/news-and-policy/bps-code-ethics-and-conduct>

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Percieved Stress. *Journal of Health and Social Behavior*, 24(4), 385–396.

Csibi, S., Griffiths, M. D., Cook, B., Demetrovics, Z., & Szabo, A. (2018). The Psychometric Properties of the Smartphone Application-Based Addiction Scale (SABAS). *International Journal of Mental Health and Addiction*, 16(2), 393–403. <https://doi.org/10.1007/s11469->

Coyne, S. M., Rogers, A. A., Zurcher, J. D., Stockdale, L., & Booth, M. (2019). Does time spent using social media impact mental health?: An eight year longitudinal study. *Computers in Human Behavior*, 104, 106160. <https://doi.org/10.1016/j.chb.2019.106160>

Cumming, G. (2014). The New Statistics: Why and How. *Psychological Science*, 25(1), 7–29. <https://doi.org/10.1177/0956797613504966>

David, M. E., Roberts, J. A., & Christenson, B. (2018). Too Much of a Good Thing: Investigating the Association between Actual Smartphone Use and Individual Well-Being. *International Journal of Human-Computer Interaction*, 34(3), 265–275. <https://doi.org/10.1080/10447318.2017.1349250>

Davidson, B. I., Shaw, H., & Ellis, D. A. (2020). Fuzzy Constructs in Assessment: The Overlap between Mental Health and Technology “Use”. *PsyArXiv*. <https://doi.org/10.31234/osf.io/6durk>

Diedenhofen, B., & Musch, J. (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, 10(4), e0121945. <https://doi.org/10.1371/journal.pone.0121945>

Elhai, J. D., & Contractor, A. A. (2018). Examining latent classes of smartphone users: Relations with psychopathology and problematic smartphone use. *Computers in Human Behavior*, 82, 159–166. <https://doi.org/10.1016/j.chb.2018.01.010>

Elhai, J. D., Dvorak, R. D., Levine, J. C., & Hall, B. J. (2017). Problematic smartphone use: A conceptual overview and systematic review of relations with anxiety and depression psychopathology. *Journal of Affective Disorders*, 207, 251–259.
<https://doi.org/10.1016/j.jad.2016.08.030>

Elhai, J. D., Yang, H., Rozgonjuk, D., & Montag, C. (2020). Using machine learning to model problematic smartphone use severity: The significant role of fear of missing out. *Addictive Behaviors*, 103, 106261. <https://doi.org/10.1016/j.addbeh.2019.106261>

Ellis, D. A. (2019). Are smartphones really that bad? Improving the psychological measurement of technology-related behaviors. *Computers in Human Behavior*, 97, 60–66.
<https://doi.org/10.1016/j.chb.2019.03.006>

Ellis, D. A. (2020). *Smartphones within Psychological Science*. Cambridge University Press

Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? *International Journal of Human-Computer Studies*, 130, 86–92.
<https://doi.org/10.1016/j.ijhcs.2019.05.004>

Funder, D. C., & Ozer, D. J. (2019). Evaluating Effect Size in Psychological Research: Sense and Nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2), 156–168.
<https://doi.org/10.1177/2515245919847202>

Fieller, E. C., Hartley, H. O., & Pearsons, E. S. (1957). Tests for rank correlation coefficients:

I. *Biometrika*, 44, 470–481.

Genc, Z. (2014). Parents' Perceptions about the Mobile Technology Use of Preschool Aged Children. *Procedia - Social and Behavioral Sciences*, 146, 55–60.

<https://doi.org/10.1016/j.sbspro.2014.08.086>

Geyer, K. (2018). *activity logger*. Retrieved from

https://play.google.com/store/apps/details?id=sensor.acitivty.android.activitylog&hl=en_GB

Geyer, K., Ellis, D. A., Shaw, H., & Davidson, B. I. (2020). Open source smartphone apps and tools for measuring, quantifying and visualizing screen time. *PsyArXiv*.

<https://doi.org/10.31234/osf.io/eqhfa>

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74–78.

<https://doi.org/10.1016/j.paid.2016.06.069>

Harris, M.M., & Bladen, A. (1994). Wording effects in the measurement of role conflict and role ambiguity: A multitrait-multimethod analysis. *Journal of Management*, 20(4), 887-901.

[https://doi.org/10.1016/0149-2063\(94\)90034-5](https://doi.org/10.1016/0149-2063(94)90034-5)

Harwood, J., Dooley, J. J., Scott, A. J., & Joiner, R. (2014). Computers in Human Behavior
Constantly connected – The effects of smart-devices on mental health. *Computers in Human Behavior*, 34, 267–272. <https://doi.org/10.1016/j.chb.2014.02.006>

Hamilton, M.T., Hamilton, D.G., & Zderic, T.W. (2007). Role of Low Energy Expenditure and Sitting in Obesity, Metabolic. *Diabetes*, 56(11), 2655–2667.

<https://doi.org/10.2337/db07-0882.CVD>

Heffer, T., Good, M., Daly, O., MacDonell, E., & Willoughby, T. (2019). The Longitudinal Association Between Social-Media Use and Depressive Symptoms Among Adolescents and Young Adults: An Empirical Reply to Twenge et al. (2018). *Clinical Psychological Science*, 1–9. <https://doi.org/10.1177/2167702618812727>

Hewson, C., Buchanan, T., Brown, I., Coulson, N., Hagger-Johnson, G., Joinson, A., ... Oates, J. (2013). Ethics Guidelines for Internet-mediated Research. In *INF206/1.2013*. Retrieved from www.bps.org.uk/publications/policy-and-guidelines/research-guidelines-policy-documents/research-guidelines-poli

Hittner, J. B., May, K., & Silver, N. C. (2003). A monte carlo evaluation of tests for comparing dependent correlations. *Journal of General Psychology*, 130(2), 149–168. <https://doi.org/10.1080/00221300309601282>

Johannes, N., Meier, A., Reinecke, L., Ehlert, S., Setiawan, D. N., Walasek, N., ... Veling, H. (2019). The Relationship Between Online Vigilance and Affective Well-Being in Everyday Life: Combining Smartphone Logging with Experience Sampling. *Preprint*. <https://doi.org/10.31234/osf.io/t3wc2>

Katevas, K., Arapakis, I., & Pielot, M. (2018). Typical Phone Use Habits: Intense Use Does Not Predict Negative Well-Being. *Proceedings of the 20th International Conference on*

Human-Computer Interaction with Mobile Devices and Services.

<https://doi.org/10.1145/3229434.3229441>

Kaye, L. K., Orben, A., Ellis, D. A., Hunter, S. C., & Houghton, S. (2020). The conceptual and methodological mayhem of “screen time.” *International Journal of Environmental Research and Public Health*, 17(10). <https://doi.org/10.3390/ijerph17103661>

Kim, J.-H. (2017). Smartphone-mediated communication vs. face-to-face interaction: Two routes to social support and problematic use of smartphone. *Computers in Human Behavior*, 67, 282–291. <https://doi.org/10.1016/j.chb.2016.11.004>

Kim, J., Hwang, Y., Kang, S., Kim, M., Kim, T.-S., Kim, J., ... Park, S. (2016). Association between Exposure to Smartphones and Ocular Health in Adolescents. *Ophthalmic Epidemiology*, 23(4), 269–276. <https://doi.org/10.3109/09286586.2015.1136652>

Kim, S.-E., Kim, J.-W., & Jee, Y.-S. (2015). Relationship between smartphone addiction and physical activity in Chinese international students in Korea. *Journal of Behavioral Addictions*, 4(3), 200–205. <https://doi.org/10.1556/2006.4.2015.028>

Kobayashi, T., & Boase, J. (2012). No Such Effect? The Implications of Measurement Error in Self-Report Measures of Mobile Communication Use. *Communication Methods and Measures*, 6(2), 126–143. <https://doi.org/10.1080/19312458.2012.679243>

Kroenke, K., Spitzer, R. L., Janet, B. W., & Williams, D. S. W. (2001). The PHQ-9: Validity of a Brief Depression Severity Measure. *Journal of General Internal Medicine*, 16(9), 606–

613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>

Kroenke, Kurt, Spitzer, R. L., Janet, B. W., Williams, D. S. ., Monahan, P., & Löwe, B. (2007). Anxiety Disorders in Primary Care: Prevalence, Impairment, Comorbidity, and Detection. *Annals of Internal Medicine*, 146, 317–325.

Kwon, M., Lee, J. Y., Won, W. Y., Park, J. W., Min, J. A., Hahn, C., ... Kim, D. J. (2013). Development and Validation of a Smartphone Addiction Scale (SAS). *PLoS ONE*, 8(2). <https://doi.org/10.1371/journal.pone.0056936>

Lee, H., Ahn, H., Nguyen, T. G., Choi, S. W., & Kim, D. J. (2017). Comparing the self-report and measured smartphone usage of college students: A pilot study. *Psychiatry Investigation*, 14(2), 198–204. <https://doi.org/10.4306/pi.2017.14.2.198>

Lepp, A., Barkley, J. E., Sanders, G. J., Rebold, M., & Gates, P. (2013). The relationship between cell phone use, physical and sedentary activity, and cardiorespiratory fitness in a sample of U.S. college students. *International Journal of Behavioral Nutrition and Physical Activity*, 10, 1–9. <https://doi.org/10.1186/1479-5868-10-79>

Loid, K., Täht, K., & Rozgonjuk, D. (2020). Do pop-up notifications regarding smartphone use decrease screen time, phone checking behavior, and self-reported problematic smartphone use? Evidence from a two-month experimental study. *Computers in Human Behavior*, 102(August 2019), 22–30. <https://doi.org/10.1016/j.chb.2019.08.007>

Meyer, G. K., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. I., Dies, R. R., ... Reed, G.

M. (2001). Psychological Testing and Psychological Assessment: A Review of Evidence and Issues. *American Psychologist Copyright*, 56(2), 128–165. <https://doi.org/10.1037//0003-066X.56.2.128>

O'Connor, S. C. & Rosenblood, L. K. (1996). Affiliation motivation in everyday experience: A theoretical comparison. *Journal of Personality and Social Psychology*, 70(3), 513–522. <https://doi.org/10.1037/0022-3514.70.3.513>

Orben, A. (2018). Cyberpsychology: A Field Lacking Theoretical Foundations. *PsyPAG The Quarterly*, (105).

Orben, A., & Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nature Human Behaviour*, 3, 173–182. <https://doi.org/10.1038/s41562-018-0506-1>

Panova, T., & Carbonell, X. (2018). Is smartphone addiction really an addiction? *Journal of Behavioral Addictions*, 7(2), 252–259. <https://doi.org/10.1556/2006.7.2018.49>

Parslow, R. C., Hepworth, S. J., & McKinney, P. A. (2003). Recall of past use of mobile phone handsets. *Radiation Protection Dosimetry*, 106(3), 233–240. <https://doi.org/10.1093/oxfordjournals.rpd.a006354>

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, 63, 539–569. <https://doi.org/10.1146/annurev-psych-120710-100452>

Regan, T., Harris, B., Van Loon, M., Nanavaty, N., Schueler, J., Engler, S., & Fields, S. A. (2020). Does mindfulness reduce the effects of risk factors for problematic smartphone use? Comparing frequency of use versus self-reported addiction. *Addictive Behaviors*, 108, 106435. <https://doi.org/10.1016/j.addbeh.2020.106435>

Richardson, M., Hussain, Z., & Griffiths, M. D. (2018). Problematic smartphone use, nature connectedness, and anxiety. *Journal of Behavioral Addictions*, 7(1), 109–116. <https://doi.org/10.1556/2006.7.2018.10>

Rozgonjuk, D., Levine, J. C., Hall, B. J., & Elhai, J. D. (2018). The association between problematic smartphone use, depression and anxiety symptom severity, and objectively measured smartphone use over one week. *Computers in Human Behavior*, 87, 10–17. <https://doi.org/10.1016/j.chb.2018.05.019>

Science and Technology Committee, UK Gov. (2019). *Impact of social media and screen-use on young people's health*. Retrieved from <https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/822/822.pdf>

Shaw, H., Ellis, D. A., & Ziegler, F. V. (2018). The Technology Integration Model (TIM). Predicting the continued use of technology. *Computers in Human Behavior*, 83, 204-214. <https://doi.org/10.1016/j.chb.2018.02.001>

Shaw, H., Ellis, D. A., Geyer, K., Ziegler, F. V. & Davidson, B. I. (2018). Quantifying smartphone ‘use’: Choice of measurement impacts relationships between ‘usage’ and health.

<https://doi.org/10.17605/OSF.IO/A4P78>

Spitzer, R. ., Kroenke, K., Janet, B. W., Williams, D., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine*, 166, 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>

Stich, J. F., Tarafdar, M., Stacey, P., & Cooper, S. C. (2019). Appraisal of email use as a source of workplace stress: A person-environment fit approach. *Journal of the Association for Information Systems*, 20(2), 132–160. <https://doi.org/10.17705/1jais.00531>

Stronge, S., Mok, T., Ejova, A., Lee, C., Zubielevitch, E., Yogeeswaran, K., ... Sibley, C. G. (2019). Social Media Use Is (Weakly) Related to Psychological Distress. *Cyberpsychology, Behavior, and Social Networking*, 22(9), 604–609. <https://doi.org/10.1089/cyber.2019.0176>

Thomée, S. (2018). Mobile Phone Use and Mental Health. A Review of the Research That Takes a Psychological Perspective on Exposure. *International Journal of Environmental Research and Public Health*, 15(12), 2692. <https://doi.org/10.3390/ijerph15122692>

Twenge, J. M. (2017). Have Smartphones destroyed a generation? Retrieved October 14, 2019, from <https://www.theatlantic.com/magazine/archive/2017/09/has-the-smartphone-destroyed-a-generation/534198/>

Twenge, J. M., Martin, G. N., & Campbell, W. K. (2018). Decreases in psychological well-being among American adolescents after 2012 and links to screen time during the rise of

smartphone technology. *Emotion*, 18(6), 765–780. <https://doi.org/10.1037/emo0000403>

Twenge, J. M. (2019). More Time on Technology, Less Happiness? Associations Between Digital-Media Use and Psychological Well-Being. *Current Directions in Psychological Science*, 28(4), 372–379. <https://doi.org/10.1177/0963721419838244>

van Smeden, M., Lash, T.L., & Groenwold, R .H. (2019). Reflection on modern methods: five myths about measurement error in epidemiological research. *International Journal of Epidemiology*, dyz251. <https://doi.org/10.1093/ije/dyz251>

Vahedi, Z., & Saiphoo, A. (2018). The association between smartphone use, stress, and anxiety: A meta-analytic review. *Stress and Health*, 34(3), 347–358. <https://doi.org/10.1002/smi.2805>

Verney, J., Schwartz, C., Amiche, S., Pereira, B., & Thivel, D. (2015). Comparisons of a Multi-Frequency Bioelectrical Impedance Analysis to the Dual-Energy X-Ray Absorptiometry Scan in Healthy Young Adults Depending on their Physical Activity Level. *Journal of Human Kinetics*, 47(1), 73–80. <https://doi.org/10.1515/hukin-2015-0063>

Vrijheid, M., Cardis, E., Armstrong, B. K., Auvinen, A., Berg, G., Blaasaas, K. G., ... Woodward, A. (2006). Validation of short term recall of mobile phone use for the Interphone study. *Occupational and Environmental Medicine*, 63(4), 237–243. <https://doi.org/10.1136/oem.2004.019281>

Wilcockson, T. D., Osborne, A. M., & Ellis, D. A. (2019). Digital detox: The effect of smartphone abstinence on mood, anxiety, and craving. *Addictive behaviors*, 99, 106013. <https://doi.org/10.1016/j.addbeh.2019.06.002>

WHO. (2018). Obesity and overweight. Retrieved September 23, 2019, from <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>

Xiang, M. Q., Lin, L., Wang, Z. R., Li, J., Xu, Z., & Hu, M. (2020). Sedentary Behavior and Problematic Smartphone Use in Chinese Adolescents: The Moderating Role of Self-Control. *Frontiers in Psychology*, 10, 1–9. <https://doi.org/10.3389/fpsyg.2019.03032>

Yang, Z., Asbury, K., & Griffiths, M. D. (2019). “A Cancer in the Minds of Youth?” A Qualitative Study of Problematic Smartphone Use among Undergraduate Students. *International Journal of Mental Health and Addiction*. <https://doi.org/10.1007/s11469-019-00204-z>

Zagalaz-Sánchez, M. L., Cachón-Zagalaz, J., Sánchez-Zafra, M., & Lara-Sánchez, A. (2019). Mini review of the use of the mobile phone and its repercussion in the deficit of physical activity. *Frontiers in Psychology*, 10, 1307. <https://doi.org/10.3389/fpsyg.2019.01307>

Zou, G. Y. (2007). Toward Using Confidence Intervals to Compare Correlations. *Psychological Methods*, 12(4), 399–413. <https://doi.org/10.1037/1082-989X.12.4.399>